Propositions, Proxies and the Paradox of Analysis

Samuel Z. Elgin*

July, 2021

Abstract

I take some initial steps toward a theory of real definition, drawing upon recent developments in higher-order logic. I demonstrate that the resulting account allows for extremely fine-grained distinctions (i.e., that it can distinguish between any relata that differ syntactically), has desirable logical features and possesses the resources to resolve the paradox of analysis.

Introduction

Definition gives rise to a number of puzzles that are not altogether easy to resolve. There is, firstly, a paradox of analysis. It is unclear how any account could be both substantive and true. For an analysis to be substantive, its content must convey information that its object does not. But if an object and content differ informationally, then the content is not the same as its object, and so the putative analysis is false. Utilitarianism is undermined neither because it ignores the distinction between persons, nor by the plethora of acts' unforeseen consequences, but rather by the mere fact that 'is morally right' differs in significance from 'maximizes utility.' And that old Aristotelian theory that to be human is, by definition, to be a rational animal is falsified neither by the theory of evolution nor by the systematic irrationality of our species, but rather by the fact that 'is human' differs in significance from 'is a rational animal.' The paradox suggests that the quest for analyses is doomed from the outset, for an account's very substance undermines its truth.

Another—and arguably intimately related—puzzle concerns the logic of definition. There are three indispensable, yet apparently incompatible, principles governing this logic: the *Identification Hypothesis*, according to which if F is, be definition, G then there is an important sense in which F is the same as G, *Leibniz's Law*, according to which terms that denote the same thing may be substituted for one another, and *Irreflexivity*, according to which there are no reflexive definitions. (As it is currently stated, the Identification Hypothesis

*I am deeply indebted to Cian Dorr—who first suggested to me that we might understand definition in terms of Fritz-style proxies.

is too imprecise to be logically respectable. This principle will be clarified throughout the course of this paper—and this explication holds the key to resolving a number of puzzles under discussion.)

Each of these principles has been assumed to be true without argument. That is to say, metaphysicians have taken each to be *so* intuitive that it is treated as a compulsory starting-point in a theory of definition.[1] But this is not to say that no defense can be given; on the contrary, there are considerations that lend support to each.

Many maintain that definition is reductive. If {Socrates} is, by definition, the set containing only Socrates then {Socrates} is nothing more than the set containing only Socrates—and if the number two is, by definition, the successor to the number one, then the number two is nothing more than the successor to the number one. Any theory of definition that denies the Identification Hypothesis appears reductive in name only (if that)—for it allows entities to be distinct from that which they are defined in terms of. And so, to account for the reductive aspect of definition, we ought to maintain that the Identification Hypothesis is true.

Leibniz's Law is also extraordinarily intuitive.[2] If Cicero is an orator and Cicero is identical to Tully, then Tully is an orator—and if Hesperus appears in the evening sky and Hesperus is identical to Phosphorus, then Phosphorus appears in the evening sky. Admittedly, there are canonical challenges for Leibniz's Law arising from opaque predicates like 'believes,' but many continue to endorse the view that identical objects bear all of the same properties—difficult cases notwithstanding.

Irreflexivity, in turn, reflects the thought that definition tracks relative fundamentality (on at least one conception of fundamentality). If water is, by definition, the chemical compound H_2O then hydrogen and oxygen are more fundamental than water is—and if hydrogen is, by definition, the element with a single proton, then protons are more fundamental than hydrogen is. A bit more precisely, we might maintain that if B occurs within the content of the definition of A, then B is more fundamental than A is.[3] Given the (not unreasonable) assumption that nothing is more fundamental than itself, it follows that there are no reflexive definitions.

But despite their initial appeal and stalwart motivation, these principles are also often rejected, for they appear to be incompatible.[4] If F is, by definition G, then—by the

[1]See, e.g., Correia (2017) for someone who endorses the Identification Hypothesis and Irreflexivity without argument, and Dorr (2016) for someone who endorses Leibniz's Law without argument.

[2]There has recently been a sustained discussion of higher-order systems that reject Leibniz's Law—see Caie, Goodman and Lederman (2020); Bacon and Russell (2019); Bacon (2019). Those who would reject Leibniz's Law in this context might appeal to the systems that they develop. I do not do so, as my account is one on which Leibniz's Law is true.

[3]See Fine (1994, 1995) for someone who defends this notion of relative fundamentality (in his words, 'ontological dependence').

[4]A bit more precisely, they are apparently incompatible on the assumption that there is at least one definition. If 'definition' had an empty extension, then the Identification Hypothesis and Irreflexivity would both be vacuously true—and there is no reason to suspect that definition would pose problems for Leibniz's

Identification Hypothesis—F is the same as G. Because F is the same as G, Leibniz's Law entails that 'F' and 'G' may be substituted for one another in every context—one of which results in 'F is, by definition, F'—contra Irreflexivity. For this reason, philosophers in this area fall into one of three camps: those who reject the Identification Hypothesis (e.g., Rosen (2015)), those who reject Leibniz's Law (e.g., Correia (2017)) and those who reject Irreflexivity (e.g., Dorr (2016)).[5] There is presently no theory of definition that validates all three principles—so it seems that any account will have an undesirable logical structure.

A third puzzle concerns the granularity of propositions and relations being defined. Arguably, definition makes extremely fine-grained distinctions. It may be that 'To be a triangle is, by definition, to be a polygon with three angles' is true while 'To be a triangle is, by definition, to be a polygon with three sides' is false. If this is so, then an adequate account of definition must distinguish between 'to be a polygon with three angles' from 'to be a polygon with three sides.' Treating properties as functions from possible worlds to extensions will not do—as the two presumably have the same extension in every possible world.

An obvious thought is to appeal to structured propositions—as they possess the requisite granularity. Perhaps propositions, properties and relations have a structure that resembles the syntactic structure of sentences and clauses; perhaps they are composed out of worldly items in much the way that sentences are composed of words.[6] In particular, it may be that *being a polygon with three angles* is composed of material concerning angles, while *being a polygon with three sides* is composed of material concerning sides. In virtue of their different composition, the two properties are distinct—their necessarily identical extensions notwithstanding—and so can stand in different relations from one another.

Accounts of structured propositions have come under sustained assault in recent years. One of their central tenets is that if the proposition Fa is identical to the proposition Gb, then $a = b$ and $F = G$. For example, if the proposition that John is a brother is identical the proposition that John is a male sibling, then John is identical to John and *being a brother* is identical to *being a male sibling.* This tenet is radically incompatible with an orthodox principle of higher-order logic: propositional identity is preserved through β-conversion.[7] Consider the binary relation R of *being the same height as*; two people stand

Law that it did not already face.

[5]Strictly, Dorr does not deny Irreflexivity here. Rather, he claims that identity performs the work of philosophical analysis, and that identity is reflexive. As I interpret Dorr (and those like him) this is as close as can be reasonably expected of a denial of Irreflexivity. Relatedly, Rosen does not explicitly deny the Identification Hypothesis, but—as Correia (2017) argues—it is possible to generate violations of Identification on his account.

[6]Rosen (2015), for example, assumes this to be true in developing his account of definition.

[7]Perhaps some balk at this logical principle precisely because it has metaphysically significant implications (in that preserving identity through β-conversion entails that the structured view of propositions is false). Those caught in the grip of a Quine's conception of logic might maintain that logic ought to be metaphysically neutral—and so any conflict between a logical system and structured propositions strikes more strongly against that system, rather than structured propositions. However, in the post-Williamson (2013)

in this relation to one another just in case they are equally tall. β-equivalence entails that $\lambda x.Rxx(a) = \lambda x.Rxa(a)$; the proposition that Mary is the same height as herself is identical to the proposition that Mary is the same height as Mary. On the structured proposition view, this entails that Mary is identical to Mary—no problems there. It also entails that $\lambda x.Rxx = \lambda x.Rxa$: the property of being the same height as oneself is identical to the property of being the same height as Mary. This is obviously absurd—the two are not even coextensive.

A second (and at least equally troubling) problem for structured propositions concerns their cardinality: the Russell-Myhill problem. Another central tenet of structured propositions is that all syntactic differences in language correspond to differences in proposition. The fact that 'Jack is to the right of Jill' differs syntactically from 'Jill is to the left of Jack' entails that the two sentences express different propositions. A bit roughly, the problem is that for every collection of propositions, it is possible to construct a sentence asserting that precisely the elements of that collection are true. For this reason, there is a mapping from the powerset of propositions to a unique sentence (i.e., there is a mapping from each element of the powerset of propositions to a sentence asserting that the propositions within that element are true). If each sentence itself corresponded to a unique proposition, then there would be a mapping from every element of the powerset of propositions to a unique proposition. But Cantor's Theorem entails that there is no such mapping. For every set s, there is no mapping from every element of the powerset of s to a unique element of s. And so, it cannot be that every syntactic difference corresponds to a propositional difference.

The upshot is this: in order for definition to make the distinctions that metaphysicians often take it to make, its relata must be finely-grained. An obvious thought is to appeal to structured propositions—as they have the requisite granularity. However, there are strong reasons to reject the structured proposition view. Those who would place a theory of definition on firm foundations ought to look elsewhere.

In light of the abundance and severity of these problems, some may be tempted to abandon the notion of definition entirely: to treat it as little more than a theoretical relic from a confused and imprecise time. I think that this is premature. Recent developments in higher-order logic suggest that the last of these puzzles, at least, admits of resolution. Fritz (Forthcoming) demonstrates that there are higher order structures that can make the fine-grained distinctions structured propositions had been intended to make, while avoiding the problems that structured propositions face. A natural thought is that definition relates terms to these higher-order proxies for structured propositions, rather than to structured propositions themselves. Interestingly, this shift resolves not only the third dilemma, but the first two as well (or so I will argue). But in order to adequately understand these resolutions, it is first necessary to describe what these proxies consist of.

era that we live in, many are susceptible to the idea that what makes something logic is its generality—not its neutrality.

Structure by Proxy

One way to frame the previous problems for structured propositions is this: it is impossible, given the proposition that Fa, to recover property F and object a, in that there may be a distinct G and b such that $Fa = Gb$. Accounts of propositions that depend upon the possibility of this recovery—like the structured proposition view—are false. Nevertheless, there is a higher-order term from which it *is* possible to recover F and a: the relation between properties and objects that has only $< F, a >$ in its extension—i.e., the relation that F stands in to a and that no other property stands in to any other object. This is not a structured proposition; after all, it isn't a proposition of any kind. It is a relation between properties and objects, and is therefore not truth-evaluable. But precisely because it is a term that allows for the recovery of F and a, it may be used as a proxy for the structured proposition that Fa—and perform some of the theoretical work metaphysicians naïvely believed structured propositions to perform.

Defining these proxies precisely requires a language to express them in. Let us adopt a simple, typed higher-order language L. Within L, there are two basic types—a type e for entities and a type t for sentences. Additionally, for any types τ_1, τ_2, $(\tau_1 \to \tau_2)$ is a type; nothing else is a type. In this language, the $\neg$ operator can be identified with a term of type $(t \to t)$. It is a function with sentences as inputs and sentences as outputs—in particular, the output is the negation of the input. Similarly, the binary operators $\wedge, \vee$ can both be identified as terms of type $(t \to (t \to t))$. Monadic first-order predicates are identified with terms of type $(e \to t)$, diadic first-order predicates are identified with terms of type $(e \to (e \to t))$, etc. There are also infinitely many variables of every type, as well as the corresponding λ abstracts needed to bind them. It would be natural, in this sort of language, to introduce quantifiers $\exists, \forall$ as terms of type $((\tau \to t) \to t)$ for every type τ. However, for the exploratory purposes of this paper I will omit any discussion of quantification.

The only additional constants worth mentioning in L are those used to express identity and definition. For every types τ_1, τ_2 there is a term $=$ of type $(\tau_1 \to (\tau_1 \to t))$, with the intended interpretation that $\ulcorner A^{\tau_1} =^{(\tau_1 \to (\tau_1 \to t)} B^{\tau_1} \urcorner$ means that A is identical to B,[8] and a term Def of type $(\tau_1 \to (\tau_2 \to t))$ with the intended interpretation that $\ulcorner Def^{(\tau_1 \to (\tau_2 \to t))}(A^{\tau_1}, B^{\tau_2}) \urcorner$ means that A is, by definition, B.[9]

There is some independent motivation for expressing a theory of definition in this sort of language. One thing such a theory might accomplish is to provide the definition of definition (or to state that there is no definition of definition, if definition is itself a primitive). Initially,

[8]Within this paper, I shift freely between infix and prefix notation.

[9]Note that, while I assume that terms must belong to the same type in order to be identical, I do not assume that terms must belong to the same type in order to be defined in terms of one another. Indeed, it is possible to make the stronger claim that, (according to the ensuing theory), there are *never* terms of the same type such that one is defined in terms of the other. I also note that, in the following I occasionally omit the types of the terms involved if it is contextually evident.

'the definition of definition' appears to assert that *definition* falls within its own extension—that is, that there is some D such that $<$ *definition*, $D >$ falls within the extension of *definition*. But if some properties or relations fall within their own extension, semantic paradox is imminent. For, it is natural to maintain that there is a property of *being a property not within its own extension*—which falls within its own extension if and only if it does not. The obvious way to accommodate this problem is to adopt a typed higher-order language like L. Rather than claiming that *definition* falls within its own extension, one may claim that a lower-order predicate 'definition' falls within the extension of a higher-order predicate 'definition.' And because the claim that a property falls within its own extension is strictly ungrammatical in these languages, paradox is avoided.

Within L, it is possibly to formally represent the aforementioned proxy for the structured proposition Fa—that is, the relation that only F stands in to a—as the bihaecceity:

$$\lambda X^{(e \to t)}.\lambda x^{e}.(X = F \wedge x = a)$$

Of course, there is nothing special about the proposition that Fa in particular; there are proxies for the structured propositions that Gb and Hc as well. It is thus desirable to provide a function that generates these proxies. This can be accomplished with the following:

$$\delta_t := \lambda X^{(e \to t)}.\lambda Y^{(e \to t)}.\lambda x^{e}.\lambda y^{e}.(X = Y \wedge x = y).$$

The δ function takes pairs of properties and objects as its input, and has—as its output—the relation that the input property stands in to the input object. This function can itself be generalized so that it provides proxies for terms of arbitrary type. For types τ_1, τ_2, there is a function:

$$\delta_{\tau_2} := \lambda X^{(\tau_1 \to \tau_2)}.\lambda Y^{(\tau_1 \to \tau_2)}.\lambda x^{\tau_1}.\lambda.y^{\tau_1}.(X = Y \wedge x = y)$$

This generates proxies for terms of type τ_2. Although this is more general than the δ_{t} function—which only generates proxies for terms of type t (indeed, it only generates proxies for a subset of terms of type t)—it still only applies to terms that stand in a particular functional relation to one another. That is, the inputs of this function must be two terms such that the latter is the functional input of the former. It will be necessary to have a function that is more general still: one whose inputs are terms of arbitrary type. For arbitrary terms τ_1, τ_2, there is a function:

$$\gamma_{\tau_1, \tau_2} := \lambda X^{\tau_1}.\lambda Y^{\tau_1}.x^{\tau_2}.y^{\tau_2}.(X = Y \wedge x = y)$$

With the γ function in place, the proxy for the proposition that Fa may be represented as:

$$\gamma(F^{e \to t}, a^e)^{10}$$

It is useful to simplify this notation still further. The result of the γ function as applied to arbitrary inputs α, β may be represented as:

$$[\alpha, \beta]$$

In particular, the proxy for the proposition that Fa is to be represented as:

$$[F, a]$$

Because the δ function is a special case of the γ function, we can represent the outputs of δ with the [ ] notation—but cannot assume that terms within brackets stand in a particular functional relation to one another without additional information about their types.

It is also valuable to define recovery functions that take proxies as their input and have, as their output, the term it is a proxy for. We may (schematically) represent these functions as:

$$Rec(\delta(\alpha, \beta)) = \alpha(\beta)^{11}$$

It may be more accurate to say that the γ function generates *immediate* proxies. However, it is desirable to also represent chains of proxies—to have a function that is sensitive not only to the immediate syntactic structure of its inputs, but the mediate syntactic structure as well. For, just as it is impossible to recover property F from proposition Fa, so too it is impossible to recover property F from the proxy $[\neg, Fa]$ (though it remains possible to recover the $\neg$ operator and proposition Fa from this proxy). In the present context, the obvious way to describe this internal structure is via recursion. Let us define a relation of decomposition, and say that one term may be decomposed into another. Decomposition is the smallest relation that satisfies the following:

1. If, $\alpha = Rec(\delta(\beta, \psi))$, then $Dec(\alpha, [\beta, \psi])$
2. If $Dec(\alpha, [\beta, \psi])$ and $Dec(\beta, [\eta, \epsilon])$ then $Dec(\alpha, [[\eta, \epsilon], \psi])$
3. If $Dec(\alpha, [\beta, \psi])$ and $Dec(\psi, [\eta, \epsilon])$ then $Dec(\alpha, [\beta, [\eta, \epsilon]])$

One notable logical feature of decomposition is that it is irreflexive and non-circular. That is to say, there are no chains of any length that take the form $Dec(\alpha, \beta), Dec(\beta, \eta)$ and $Dec(\eta, \alpha)$.[12]

[10] Here, as elsewhere, I assume that identity is preserved through $\beta\eta$ conversion—so that $\lambda X^{(e \to t)}.\lambda Y^{(e \to t)}.\lambda x^e.\lambda.y^e.(X = Y \wedge x = y)(F, a)$ is identical to $\lambda X^{(e \to t)}.\lambda x^e.(X = F \wedge x = a)$.

[11] Note that the *Rec* function takes applications of the δ function—rather than the γ function—as its inputs, so we may indeed assume that the latter term is the functional input of the former.

[12] PROOF: of irreflexivity:

A term might have any number of decompositions; $\neg Fa$ may be decomposed as

$$[\neg, Fa]$$

or, alternatively, as

$$[\neg, [F, a]]^{13}$$

On one end of the spectrum, decomposition is sensitive only to the outermost syntactic structure of a term, while—at the other end of the spectrum—a decomposition reveals the entirety of a term's syntactic structure. The latter of these is of particular interest, so it is valuable to define notation for that instance of decomposition. This notation need not be taken as primitive; it may be introduced recursively based upon the notation already to hand (defining it only for terms involving constants in function-application positions):

$\lceil AB \rceil = [A, B]$	for constants $A^{\tau_1 \to \tau_2}, B^{\tau_1}$
$\lceil A\beta \rceil = [A, \lceil \beta \rceil]$	for constant $A^{\tau_1 \to \tau_2}$ and non-constant β
$\lceil \alpha B \rceil = [\lceil \alpha \rceil, B]$	for non-constant α and constant B^{τ_1}
$\lceil \alpha\beta \rceil = [\lceil \alpha \rceil, \lceil \beta \rceil]$	for non-constants α, β

As it has been defined, there is a connection between the $\lceil\ \rceil$ function and decomposition;

Define a function *Rank* from types to the natural numbers, such that $Rank(e) = 1$ and $Rank(t) = 1$ and, for types τ_1, τ_2, $Rank(\tau_1 \to \tau_2) = Rank(\tau_1) + Rank(\tau_2)$. Trivially every type has the same rank as itself, so—to prove irreflexivity—it suffices to show that if $Dec(A^{\tau_1}, B^{\tau_2})$, then $Rank(\tau_1) \neq Rank(\tau_2)$. In particular, it will be shown that if $Dec(A^{\tau_1}, B^{\tau_2})$, then $Rank(\tau_1) < Rank(\tau_2)$.

Base Case: Suppose $\alpha = Rec(\delta(\beta, \psi))$. Then $\alpha = \beta(\psi)$. So, if β is of type $(\tau_1 \to \tau_2)$ and ψ is of type τ_1 then α is of type τ_2. The type of $[\beta, \psi]$ is $((\tau_1 \to \tau_2) \to (\tau_1 \to t))$. $Rank(((\tau_1 \to \tau_2) \to (\tau_1 \to t))) = 2Rank(\tau_1) + Rank(\tau_2) + 1 > Rank(\tau_2)$.

Inductive Step: Suppose $Dec(\alpha^{\tau_1}, [\beta^{\tau_2}, \psi^{\tau_3}])$ and $Dec(\beta^{\tau_2}, [\eta^{\tau_4}, \epsilon^{\tau_5}])$ such that $Rank(\tau_1) < Rank(\tau_2 \to (\tau_3 \to t))$ and $Rank(\tau_2) < Rank(\tau_4 \to (\tau_5 \to t))$. Let

$Rank(\tau_1) = n$
$Rank(\tau_2) = m$
$Rank(\tau_3) = o$
$Rank(\tau_4) = p$
$Rank(\tau_5) = q$

Therefore, $n < m + o + 1$ and $m < p + q + 1$. The type of $[[\eta, \epsilon], \psi]$ is $(((\tau_4 \to (\tau_5 \to t)) \to \tau_3) \to t)$. The rank of this type is $p + q + o + 2$. $p + q + o + 2 > m + o + 1 > n$. Therefore, $Rank(\tau_1) < Rank(((\tau_4 \to (\tau_5 \to t)) \to \tau_3) \to t)$.

The second inductive step can be demonstrated analogously to the first.

Therefore, Decomposition is irreflexive. Given the directionality of *Rank*, it should be clear how this also precludes decompositional circles.

[13] This example underscores the need for the γ function. Had we operated only with the δ function, the expression $[\neg, [F, a]]$ would be ungrammatical, as the $[\ ]$ notation would require that the second term—$[F, a]$—be a functional input of the former term—$\neg$ (which it isn't). By shifting to the γ function, we may express this type of proxy grammatically.

that is to say, if $\lceil \alpha \rceil = \beta$, then $Dec(\alpha, \beta)$.[14] This notation represents the ultimate proxy for the structured proposition that $\neg Fa$ as:

$$\lceil \neg Fa \rceil = [\neg, \lceil Fa \rceil] = [\neg, [F, a]]$$

Another, only slightly more complex, example is:

$$\lceil Fa \wedge Gb \rceil = [\lceil Fa \rceil, \lceil \wedge Gb \rceil] = [[F, a], [\wedge, \lceil Gb \rceil]] = [[F, a], [\wedge, [G, b]]]$$

Note that the $\lceil\ \rceil$ notation is such that the types of these terms depend not merely upon the types of the terms occurring within its scope, but to their syntactic structure as well; while $\lceil \neg\neg Fa \rceil$ strongly resembles $\lceil Fa \rceil$ the terms are of entirely different types.

These formalisms suffice for the purposes of this paper. As is doubtlessly already clear, proxies make some of the same distinctions that structured propositions had been intended to make; for any α and β that differ syntactically, $\lceil \alpha \rceil$ is distinct from $\lceil \beta \rceil$. But we can be relatively confident that the theory of proxies is not inconsistent as the theory of structured propositions is. There are no resources within this language that surpass the descriptive power of traditional higher-order languages with lambda abstraction and terms for identity. Rather, we have merely introduced shorthands to identify particular relations within this existing framework. This framework itself has been proven to be consistent—so the risk that these proxies engender contradiction is minimal.[15]

It is a matter of some interest why the theory of proxies is consistent while the theory of structured propositions is not. The Russell-Myhill problem depended upon the ability to associate a structured proposition with every collection of propositions. That is, for every collection of propositions, it was possible to generate a proposition asserting that all and only the elements of that collection were true. In the limiting case, there is a proposition

[14] Proof by induction:

Base case: Suppose that $A^{\tau_1 \to \tau_2}$ and B^{τ_1} are constants. Then $\lceil AB \rceil = [A, B]$. From the definition of decomposition, it follows that $Dec(AB, [A, B])$.

First inductive step: Select an $A\beta$ such that A is a constant and β is not a constant such that if $\lceil \beta \rceil = \eta$ then $Dec(\beta, \eta)$. Show that if $\lceil A\beta \rceil = \psi$ then $Dec(A\beta, \psi)$.

Select an arbitrary η such that $\lceil \beta \rceil = \eta$.

Recall that $\lceil\ \rceil$ is only defined over constants in function application position. Therefore, it must be that β is a functional input of A. From the definition of decomposition, we have $Dec(A\beta, [A, \beta])$.

From the definition of Dec, it follows that $Dec(A\beta, [A, \eta])$.

$\lceil A\beta \rceil = [A, \lceil \beta \rceil] = [A, \eta]$.

Therefore, if $\lceil A\beta \rceil = \psi$ then $Dec(A\beta, \psi)$

The second and third inductive steps are proven analogously to the first.

Therefore, if $\lceil \alpha \rceil = \beta$ then $Dec(\alpha, \beta)$.

[15] The reason I say that we can be 'relatively confident' in the consistency of this language and that there is a 'minimal risk' of contradiction is that—as per Gödel's second incompleteness theorem—the systems used to prove the consistency of this language are more expressively powerful (and thus more likely to lead to contradiction) than language L itself. Suffice it to say that no one has yet uncovered an inconsistency in this type of language and that—if there is one—the problems that that would generate would far exceed this paper.

asserting (albeit falsely) that all propositions are true. Proxies perform some of the theoretical functions of structured propositions. In order to generate an analogous problem for proxies, we would require the ability to make a claim about all proxies of arbitrary type. And this is something that the present theory can*not* do. Language L cannot consistently have sufficiently many propositions that we can assign a unique proposition to each collection of proxies. That is to say, while it is possible to generate propositions that make claims about all proxies of type τ (for an arbitrary τ), it is impossible to generalize and thereby make a claim about proxies of any type. This is the cost of consistency—it is a price I will pay.

Definition by Proxy

Recall that definition gave rise to a number of puzzles, and that I have suggested that the relation of decomposition holds the key to resolving them all. Of course, decomposition itself has no immediate implications for definition; some principle is required linking the two. For the remainder of this paper, I will explore—and to some extent defend—the following:[16]

> **Definition by Proxy (DBP)**
> If $Def(\alpha, \beta)$, then $Dec(\alpha, \beta)$

I.e., if α is, by definition, β, then β is a decomposition of α.

There are several ways that this account might be strengthened. The conditional might, firstly, be replaced with a biconditional—so that $Def(\alpha, \beta)$ holds just in case $Dec(\alpha, \beta)$. DBP states merely that all definitions are decompositions, but this second principle also holds that all decompositions are themselves definitions. And so, while one can consistently hold that DBP is true and $Def(Fa, [F, a])$, is false, this is not so for this stronger principle; the fact that Fa is decomposed into $[F, a]$ guarantees that the proposition is defined in terms of that proxy. This biconditional might itself be strengthened still further to constitute a definition of definition. That is, not only are definition and decomposition coextensive, but we can define definition itself in terms of the relation of decomposition.

There are several ways to characterize this type of account. In some respects, it might be considered hylomorphic—where definition is to be understood in terms of matter and form. The 'matter' for the definitions of properties, propositions and relations are given by the constants in our language L, and the 'form' is determined by their logical form. If, for example, $Def(Fa, [F, a])$ then the proposition that Fa is defined in terms of property F

[16] The reason I say 'to some extent' is that what I immediately aim to demonstrate is that this principle resolves the puzzles mentioned at the outset of this paper. I do not contrast this account with other conceptions of definition—nor do I claim that it satisfies any other theoretical requirements that definition ought to satisfy. Nevertheless, I take it that these problems are intractable enough that the ability to resolve them counts substantially in this principle's favor.

and object a, and the form of this definition is determined by the logical types of F $(e \to t)$ and a (e).

Another method of strengthening Definition by Proxy involves identifying a unique proxy relevant to a definition. For, as we have already seen, a term might have any number of decompositions. Those who maintain that each term has a single definition thus cannot accept that all decompositions are definitions—a unique proxy must be identified. Along these lines, it is natural to hold that the relevant proxy is given by the $\lceil\ \rceil$ function—as this provides the maximal information (in that it is possible to extract the most terms from this proxy). Definition by Proxy may thus become $Def(\alpha, \beta) \to Dec(\alpha, \lceil\eta\rceil)$ (where $\beta = \lceil\eta\rceil$). But the claim that each term has a unique definition is (or, at least, ought to be) controversial. Many hold that definition is transitive; if A is, by definition, B and B is, by definition, C then A is, by definition, C. Those who accept that transitivity holds (nonvacuously) cannot also hold that each term has a unique definition.

Each of these strengthened principles is independently worthy of consideration, but I will say no more about them here. It is my aim to resolve the puzzles mentioned at the outset of this paper and, to that end, DBP will suffice.

The Granularity of Definition

There was a puzzle concerning the granularity of definition—it was unclear how definition could make the distinctions that metaphysicians typically take it to make. Quite plausibly, 'To be a triangle is, by definition to be a polygon with three angles' is true while 'To be a triangle is, by definition, to be a polygon with three sides' is false—despite the fact that any object bearing one property necessarily bears the other. If so, an account of definition must be capable of distinguishing 'To be a polygon with three angles' from 'To be a polygon with three sides'—an extremely fine-grained distinction.

DBP is capable of making this type of distinction. If terms α, β within L differ syntactically, $\lceil\alpha\rceil$ denotes a different proxy than $\lceil\beta\rceil$ denotes. For example, $\lceil p \wedge q\rceil$ differs in denotation from $\lceil q \wedge p\rceil$ (on the assumption that p and q are distinct); while the first refers to the relation that p stands in to (the relation that $\wedge$ stands into q), the second denotes the relation that q stands in to (the relation that $\wedge$ stands into p). Because definition entails decomposition, it may be that $Def(p \wedge q, \lceil p \wedge q\rceil)$ is true while $Def(p \wedge q, \lceil q \wedge p\rceil)$ is false; the conjunction is defined in terms of one proxy but not the other.

The distinction between 'To be a polygon with three angles' and 'To be a polygon with three sides' is conceptually analogous, but the details are slightly more cumbersome due to the presence of 'three.' On a roughly Fregean approach, this term refers to the higher-order property of 'is a property with three objects within its extension.' Technicalities aside, the proxy which relates this property to 'is a side' differs from the proxy that relates this to 'is an angle.' And, for this reason, it may be that 'To be a triangle is, by definition, to be a polygon with three sides' is true while 'To be a triangle is, by definition, to be a polygon with three angles' is false. DBP is thus capable of making extraordinarily fine-grained distinctions—indeed, as finely grained as a metaphysician could possibly require.

The Logic of Definition

There were three principles regarding the logic of definition that were exceedingly natural, yet are apparently in conflict. The principles are:

The Identification Hypothesis	If F is, by definition, G then there is an important sense in which F is the same as G.
Leibniz's Law	Terms that denote the same thing may be substituted for one another.
Irreflexivity	There are no reflexive definitions.

The easiest of these to accommodate is *irreflexivity.* According to DBP, if α is, by definition, β, then β is a decomposition of α. Because decomposition is irreflexive, it follows that definition is irreflexive; there are no true instances of $Def(\alpha, \alpha)$.

It is less obvious how DBP accommodates the Identification Hypothesis—at least partially due to the imprecision with which it is stated. Because definition is irreflexive while identity is reflexive, it cannot be that terms are literally identical to that which they are defined in terms of. Nevertheless, there remains an important connections between definition and identity that results from this account:[17]

Given an arbitrary α, that lacks the [ ] notation:
$$Def(\alpha, \lceil\beta\rceil) \rightarrow \alpha = \beta$$

[17]LEMMA: If α, β are terms that lack the [ ] notation, then $Dec(\alpha, \eta) \wedge Dec(\beta, \eta) \rightarrow \alpha = \beta$.
PROOF: by induction on Dec.
Base Case: Trivially, if $\alpha = Rec(\delta(\phi, \psi))$ and $\beta = Rec(\delta(\phi, \psi))$ then $Dec(\alpha, [\phi, \psi])$ and $Dec(\beta, [\phi, \psi])$ and $\alpha = \beta$.
Inductive Case: Suppose that ϕ, ψ have the inductive property.
Suppose $Dec(\alpha, [\phi, \psi])$
Suppose $Dec(\beta, [\phi, \psi])$
Because ϕ and ψ have the inductive property, there must exist a unique ω, ϵ (which may or may not be identical to ϕ and ψ) that lack the [ ] notation such that $Dec(\alpha, [\omega, \epsilon])$ and $Dec(\beta, [\omega, \epsilon])$. That is, if ϕ lacks the [ ] notation then we may let $\omega = \phi$. If ϕ has the [ ] notation then it is a proxy for some term that lacks the [ ] notation. Given the inductive hypothesis, this term is unique—which we denote with ω. Just so for ψ and ϵ.
Because ω and ϵ lack the [ ] notation, it must be that $\alpha = Rec(\delta(\omega, \epsilon)) = \beta$.
Therefore, if α, β are terms that lack the [ ] notation, then $Dec(\alpha, \eta) \wedge Dec(\beta, \eta) \rightarrow \alpha = \beta$.

PROVE: $Def(\alpha, \lceil\beta\rceil) \rightarrow \alpha = \beta$.
Suppose $Def(\alpha, \lceil\beta\rceil)$. Given DBP, it follows that $Dec(\alpha, \lceil\beta\rceil)$.
Given the proof in footnote 14, it follows that $Dec(\beta, \lceil\beta\rceil)$.
We have $Dec(\alpha, \lceil\beta\rceil)$ and $Dec(\beta, \lceil\beta\rceil)$. Given Lemma, it follows that $\alpha = \beta$.

This claims that, for an arbitrary term α that is not itself a proxy, if α is defined in terms of the final proxy for β, then α is identical to β. For example, if $Def(Fa, \lceil Gb \rceil)$, then $Fa = Gb$. That is to say, if α is, by definition, β, then α is identical to that which β is a proxy for. Because of this connection, the Identification Hypothesis is true.

It less obvious still how this accommodates Leibniz's Law. DBP has implications that initially appear to ensure opacity. For example, it may be that both $Fa = Gb$ and $Def(Fa, \lceil Fa \rceil)$ are true while $Def(Fa, \lceil Gb \rceil)$ is false. For this reason, some might reasonably suspect that 'Fa' cannot be substituted for 'Gb' in every context—despite the fact that the two denote the same proposition.

This is not actually a violation of Leibniz's Law. Leibniz's Law asserts that terms that denote the same object may be substituted for one another *salve veritate*. But when the term 'Fa' appears within the $\lceil\ \rceil$ notation, it does not actually refer to the proposition Fa—rather, it refers to the relation that F stands in to a. And although Fa is identical to Gb, $\lceil Fa \rceil$ is not identical to $\lceil Gb \rceil$ if F is not identical to G. Because the terms $\lceil Fa \rceil$ and $\lceil Gb \rceil$ have different referents, Leibniz's Law is inapplicable.

To alleviate this concern, it may help to draw an analogy to quotation. Although 'Hesperus' and 'Phosphorus' both refer to Venus, the terms may not be substituted for one another when appearing within quotation marks. While Hesperus is Phosphorus and 'Hesperus' begins with the letter H, Leibniz's Law does not entail that 'Phosphorus' does as well. Within quotation, the terms refer to the words themselves rather than to their ordinary referents—and identity of referents does not guarantee the identity of words that denote them. In this respect, the $\lceil\ \rceil$ notation functions like quotation; a term has a different referent within this notation than it has out of it. While 'Fa' typically refers to the proposition that Fa, when it appears within $\lceil\ \rceil$ it refers to the relation F stands in to a. If we would like, we could eliminate the $\lceil\ \rceil$ notation from our language, and represent proxies solely with the $[\ ]$ notation. Once this is done, the previous example becomes $Fa = Gb$ and $Def(Fa, [F, a])$ are true while $Def(Fa, [G, b])$ is false—which is no violation of Leibniz's Law.

The upshot is this: DBP entails Irreflexivity and the Identification Hypothesis. And while it does not itself guarantee the truth of Leibniz's Law, the apparent conflict between the two is illusory; it thus provides no reason to reject Leibniz's Law beyond the puzzles it already faced. The ability to reconcile three seemingly incompatible principles counts in its favor.

The Paradox of Analysis

The paradox of analysis concerns the ability of substantive analyses to be true. For an analysis to be substantive, its content must possess information that its object does not. But if an object and content differed informationally, there seemed no way for the content to be the same as its object—and so the putative analysis is false.

It is my hope that the development of DBP indicates how the paradox is to be resolved.

There is an important sense in which every analysis is substantive; information is present in the content of an analysis that is missing from the object. If $Def(Fa, [F, a])$, then the content of analysis is $[F, a]$—a term from which F may be extracted, while (as already belabored at length) it is impossible to extract F from Fa. Similar results hold for all definitions; if definition entails decomposition, then terms can be extracted from the content of a definition that cannot be extracted from the object.

This difference in substance poses no threat to the Identification Hypothesis. Although information can be extracted form a proxy that cannot be extracted from the term it is a proxy for, this does not falsify the claim that the term it is a proxy for is identical to the object of analysis. DBP thus provides a path to substantive definition.

There is another respect in which definitions can be informative. β may be a proxy for α while agents are unaware that β is a proxy for α; perhaps some are ignorant that Fa is decomposed in the matter that it is. If definition by proxy is true, then definition reveals (at least one of) the proxies for a term. And so, by coming to recognize facts about definition, agents may thereby come to recognize facts about proxies.

Conclusion

Definition by Proxy thus resolves the puzzles mentioned at the outset of this paper. Definition is capable of distinguishing relata as finely grained as the syntax of language L. It entails that the Identification Hypothesis and Irreflexivity are true—and these do not appear to conflict with Leibniz's Law. And, because proxies possess informational content that the terms they are proxies for lack, definitions are substantive.

I close by addressing a potential objection to this theory that was mentioned to me by Peter Fritz. Many believe that the paradox of analysis is really nothing other than an iteration of the Frege/Mates puzzle of identity and synonymy. Definitions are substantive in the way that other identity claims may be substantive. While this thought does not itself resolve the paradox of analysis (in that it does not provide a resolution to the Frege/Mates puzzle), it suggests that any resolution of the paradox of analysis ought to be applicable to the general puzzle of identity as well. And, precisely because the current proposed solution has no obvious implications for Frege/Mates, it is inadequate.

I am tempted to turn this objection on its head. Regardless of whether my proposed account of definition is to be accepted, it is at least a *potential* path toward resolving the paradox of analysis. And, precisely because there is a potential resolution to the paradox of analysis that is not a resolution to Frege/Mates, the paradox of analysis is distinct from the Frege/Mates puzzle.

References

Bacon, Andrew. 2019. "Substitution Structures." *Journal of Philosophical Logic* 48:1017–75.

Bacon, Andrew and Jeffrey Russell. 2019. "The Logic of Opacity." *Philosophy and Phenomenological Research* 99(1):81–114.

Caie, Michael, Jeremy Goodman and Harvey Lederman. 2020. "Classical Opacity." *Philosophy and Phenomenological Research* 101(3):524–66.

Correia, Fabrice. 2017. "Real Definitions." *Philosophical Issues* 27(1):52–73.

Dorr, Cian. 2016. "To be F is to be G." *Philosophical Perspectives* 30(1):39–134.

Fine, Kit. 1994. "Essence and Modality." *Philosophical Perspectives* 8:1–16.

Fine, Kit. 1995. "Ontological Dependence." *Proceedings of the Aristotelian Society* 95:269–90.

Fritz, Peter. Forthcoming. "Structure by Proxy with an Application to Grounding." *Synthese* .

Rosen, Gideon. 2015. "Real Definition." *Analytic Philosophy* 56(3):189–205.

Williamson, Timothy. 2013. *Modal Logic as Metaphysics.* Oxford University Press.